

PAPER • OPEN ACCESS

Flower Recognition using Deep Convolutional Neural Networks

To cite this article: Mastura Hanafiah *et al* 2022 *IOP Conf. Ser.: Earth Environ. Sci.* **1019** 012021

View the [article online](#) for updates and enhancements.

You may also like

- [Detection of passageways in natural foliage using biomimetic sonar](#)
Ruihao Wang, Yimeng Liu and Rolf Müller
- [Vegetable Pest Image Recognition Method Based on Improved VGG Convolution Neural Network](#)
Haijian Ye, Hang Han, Linna Zhu et al.
- [A deep convolutional neural network model for hand gesture recognition in 2D near-infrared images](#)
Celal Can, Yasin Kaya and Fatih Kılıç

ECS Toyota Young Investigator Fellowship

For young professionals and scholars pursuing research in batteries, fuel cells and hydrogen, and future sustainable technologies.

At least one \$50,000 fellowship is available annually.
More than \$1.4 million awarded since 2015!



Application deadline: January 31, 2023



TOYOTA

Learn more. Apply today!

Flower Recognition using Deep Convolutional Neural Networks

Mastura Hanafiah¹, Mohd Azraei Adnan¹, Shuzlina Abdul-Rahman^{1,3}, Sofianita Motalib^{1,3}, Ariff Md Ab Malik^{2,3}, Mohd Razif Shamsuddin^{1,3}

¹ Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Shah Alam, Malaysia

² Faculty of Business and Management, Puncak Alam Campus, Universiti Teknologi MARA Selangor, Puncak Alam, Selangor, Malaysia

³ Research Initiative Group of Intelligent Systems, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Shah Alam, Malaysia

Corresponding author's email: mastura.hanafiah@uitm.edu.my

Abstract. This study investigates the suitable model for flower recognition based on deep *Convolutional Neural Networks* (CNN) with transfer learning approach. The dataset used in the study is a benchmark dataset from Kaggle. The performance of CNN for plant identification using images of flower are investigated using two popular image classification models: *AlexNet* and *VGG16*. Results show that CNN is proven to produce outstanding results for object recognition, but its achievement can still be influenced by the type of images and the number of layers of the CNN architecture. The models produced adequate performance rates, with the *VGG16* model achieving the best results. *AlexNet* and *VGG16* models achieved the accuracy of 85.69% and 95.02% respectively. This model can be replicated for flower recognition in other areas, especially in our national heritage, Taman Negara which is among the richest flora ecosystem in the world. The significant feature extraction processes were discussed as well, and this is useful for other types of flowers than the trained dataset.

Keywords: Flower Recognition, Convolutional Neural Networks, SDG, Taman Negara Pahang

1. Introduction

The essence of flowers has been extracted and manipulated in both traditional and modern medicine, and flowers are found with different species and special features. The recognition of flower species requires significant knowledge from the field of botanical science and experiences. With the growth of machine learning, recognition algorithms have indeed aided scientists to identify plants including flowers for further scientific research [1, 2]. Recognition of numerous plant species will support the development of the pharmaceutical industry especially now as medical plants are considered to having lesser adverse reactions and being less expensive than contemporary medication [3]. Flower classification or recognition via machine learning has evolved since the last decade. Methods such as *Support Vector Machine* (SVM) [4], *Markov Random Field* (MRF) [5], and *Back Propagation Artificial*



Neural Network (ANN) [6] have been used. Accuracies of 86.17% with Oxford 17 dataset [4], 94% with Oxford 17 dataset [5], and 81.19% with Oxford 102 dataset [6] had been achieved respectively.

While flower classification is tempting in terms of use and meaning, it has been hindered by a number of constraints. Due to the many varieties of flowers, the classification of flowers has been more complex. Flower objects are not like any other obvious objects where the categories can be distinguished easily. Flower classifications are more difficult to perform due to the inter-class similarity and substantial intra-class variation issues [7]. Inter-class similarity refers to the flowers of many types yet with a similar appearance, while intra-class variation is caused by illumination changes and view differences. In another study, a method that needs the user to put a cloth of black colour behind the flower object was used in order to distinguish flowers [8], however, this method is both infeasible and impractical. These complexities have caused confusion among flower classes, making flower classification even more difficult.

Recent development in computer vision has shown achievement in feature representation with deep learning *Convolutional Neural Network* (CNN) such as object detection, segmentation, and image classification [9, 10]. Thus, the aim of this study is to investigate the suitable model based on deep learning CNN approach that will be able to classify and distinguish the various species of flower.

2. Related works

Retrospectively, with an estimated age of at least 130 million years, Malaysia National Park is in a class of its own as Malaysia's oldest and most popular national park [11]. The park is home to indigenous people with plenty of flora and animals to keep researchers interested in its natural history. Additionally, this national park boasts 10,000 species of plants; an astounding 150,000 varieties of insects [12]. The rich content of this park is one of many developments that have taken place over the last decades and has contributed to the Malaysian tourism economy. The Malaysia Government programs have encouraged many research and technology developments to explore and preserve this tropical and national heritage. One of the past studies was conducted by Suratman at Kuala Keniam, Pahang National Park [13]. According to the findings, the woods of Kuala Keniam are characterised by a uniform distribution of individuals with mixed species composition, with varying combinations of dominant and codominant species representing distinct sites. The findings from Suratman [13] show that many more studies can be conducted in exploiting the richness of species of Kuala Keniam. There are still abundant names and varieties of wildflowers that we do not recognize.

2.1. Flower Recognition

The process of quantifying flowers normally involves three important features: color, texture, and shape [6, 14]. *Hue, Saturation, and Value* (HSV) is used to describe colors, as compared to *Red, Green, Blue* (RGB), as it can eliminate the false data caused by variation lighting condition. The hue or (tint) describes the shade of the colors, the saturation describes the amount of gray level and the value describes the brightness or luminance [6]. Gray Level Co-occurrence Matrices (GLCM) were utilized by Haralick texture analysis to specify 14 statistical features that could be computed to quantify an image based on texture by lowering the image grey levels [15]. These texture features are widely used in a variety of texture analysis projects due to their simplicity and intuitive interpretations [16]. Moment invariants have been extensively used in shape recognition studies. The two most often used global form descriptors are *Hu moments* and *Zernike moments*. In *Hu moments*, there are seven invariant moments: the mean, variance, standard deviation, skew, kurtosis, and other statistical parameters [17]. These seven moments are combined to generate a 7-dimensional feature vector. Whereas for *Zernike moment*, rotation and scale invariance, resilience to noise, expression efficiency, and quick calculation are all desirable moment descriptors [14].

2.2. Image Classification via Machine Learning

Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Multi-Layer Perceptron (MLP), and Convolutional Neural Networks (CNN) are examples of machine learning (ML) techniques that can be used to classify images. SVM defines decision boundaries with hyperplanes that separate two classes of data from each other and put them into a feature space. Although SVM can perform a non-linear classification, it has some restrictions in terms of speed and size during training and testing phases of the algorithm [18]. For KNN, although it is simple and multi-class capable, it employs all features equally for similarity calculating, which might result in classification errors [18]. ML has evolved into deep learning in which MLP and CNN are the most common algorithms that are based on *Artificial Neural Network* (ANN). Between the input and output layers, MLP employs a hidden layer. It can learn non-linear models, but it has drawbacks including a non-convex loss function and tuning for the number of hidden layers and neurons [19]. With increased processing power, CNN, on the other hand, has been able to achieve superhuman performance on some challenging visual tasks [20].

2.3. Convolutional Neural Network

CNN involves image classification and object recognition, and thus, it is suitable to be employed in this study as the flower recognition system must detect the flower image and further classifies which flower species it belongs to [14]. CNN consists of five layers: 2 convolutional layers, 2 pooling layers and 1 fully connected layer, as shown in figure 1. The convolutional layer convolutes the input images with learnable filters and extracts the features to generate feature maps. The maps become more insensitive to rotation and distortion towards the higher layer by providing more and more complex generalizations. The pooling layer performs spatial dimensionality reduction on the feature maps to only keep the important features. One or more fully connected layers receive the final output from the convolutional and pooling layers. The output prediction is then obtained by passing the data to the classifier layer, which employs activation functions such as Softmax.

The first convolutional layer in the input layer is set to 224 x 224 RGB images. The image is passed through a series of convolutional (conv.) layers, each with a very small receptive field: 3×3 pixels (which is the smallest size to capture the notion of left, right, up, down, center). The Fully Connected (FC) layer of the output layer is fully connected to the neurons in the previous layer.

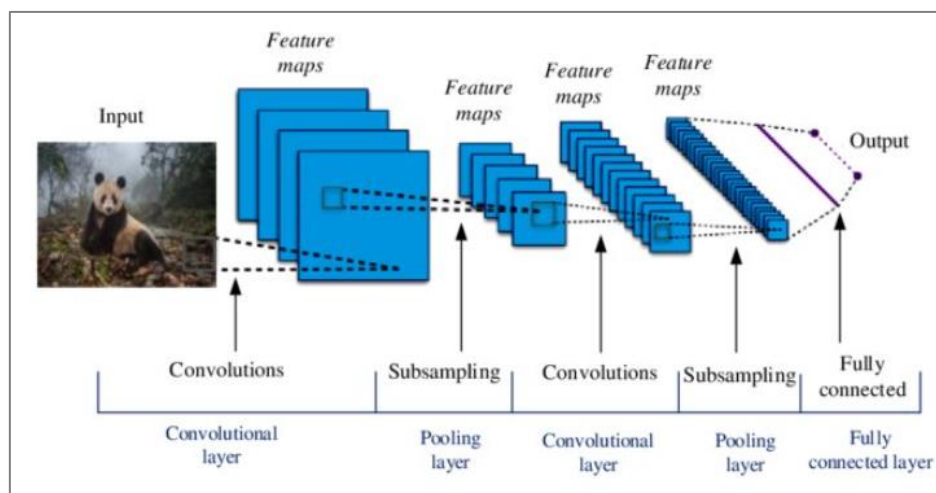


Figure 1. A simple CNN architecture [24].

The FC layer predicts the input image's final class or label in which it produces its output with dimension $[1 \times 1 \times N]$, with N denoting the number of classes or labels t to be considered for the classification. The classification of the flower species is achieved by using transfer learning method from an already pre-trained network. Transfer learning for CNN is the process of learning the training

model on a specific data set and then applying the parameters to the new domain's target dataset. Not only it can learn colors, texture, and other low-level features from the training dataset, but it can also learn the complex semantic features to aid in the categorization of the target dataset, improving the model's classification performance [21].

A flower species detection system based on CNN and transfer learning was developed, with a classifier like logistic regression or random forest applied on top to improve accuracy [14]. As a feature extractor, CNN paired with transfer learning beat all other handmade feature extraction approaches, according to the findings. Accuracies of 73.05%, 93.41%, and 90.60 were achieved using feature extractors which are *LearningOverFeat*, *Inception-v3*, and *Xception* architectures respectively. CNN with transfer learning also yielded outstanding outcomes for similar flower classification systems [21]. Oxford flower dataset was used, and *VGG-16*, *VGG-19*, *Inception-v3*, and *ResNet50* models were used to compare the network initialization model with the transfer learning approach. The findings suggest that transfer learning can successfully avoid deep convolutional networks, which are prone to local optimum and over-fitting issues. It is also demonstrated that *Inception-v3-transfer* and *ResNet50-transfer* have higher classification accuracy than the other approaches.

3. Methodology

3.1. Research flow

Figure 2 below shows the overall research flow. The process started with loading the images into the model. Image pre-processing were required before the images could be used as a training dataset. The dataset was split into the ratio of 70:30 to avoid overfitting where 70% was used for the training dataset, while another 30% was used as the testing dataset. Additionally, this division was meant to avoid overestimation of accuracy. The approach would produce the training and testing dataset's experimental outcome. The activity continued with an evaluation and analysis of the results generated by the preceding phase in order to determine the model's accuracy.

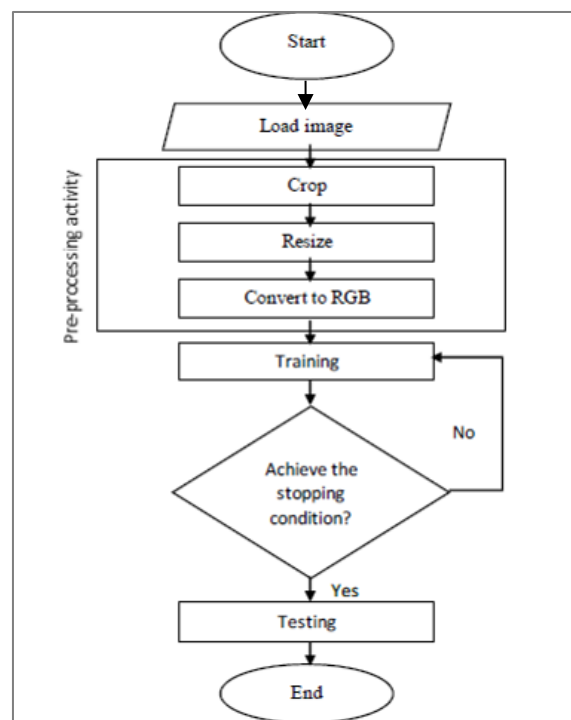


Figure 2. Research flow.

3.2. Image collection and pre-processing

The next stage was data acquisition or image collection. Due to limited samples of flowers during our trip in Taman Negara, the benchmark dataset was retrieved from the Kaggle website (<https://www.kaggle.com/alxmamaev/flowers-recognition>). A total of 2,648 flower images had been collected consisting of five different flower classes which were daisy, dandelion, rose, sunflower, and tulip. Each category contained between 438 and 709 images. Only images captured in the daylight were selected as the illumination for the night mode would be different. A mixed view angle was also covered.

Because some flower categories had very distinct and distinguishing qualities, it was difficult to differentiate some of them because they were so similar, such as dandelion and sunflower. This dataset was also collected at various scales and illumination levels. It was more difficult to classify due to the diversity of the classes and the slight distinctions between each category. Figure 3 shows the classes of flowers in the scope of the study with some sample images.

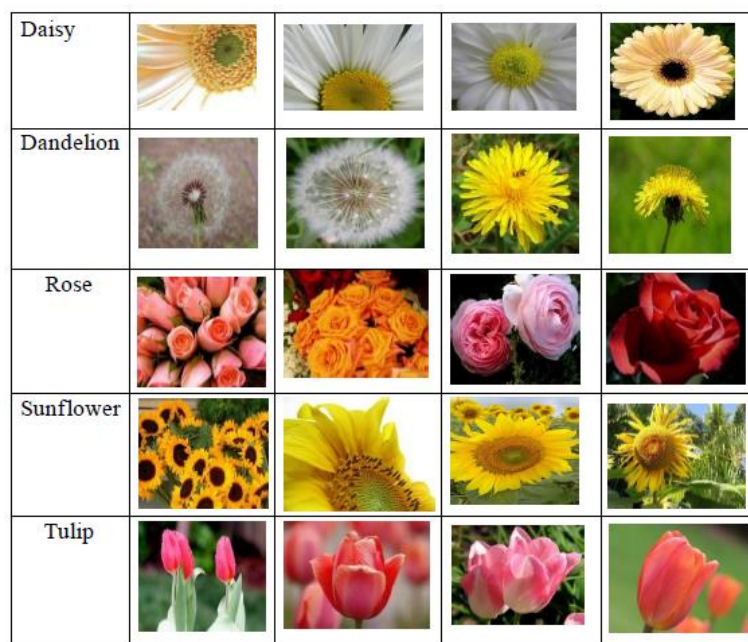


Figure 3. Sample of a dataset for five species.

The images were required to be pre-processed before they can be used as the training dataset. *Python programming* with *OpenCV*, and *matplotlib* libraries were used for these activities. Images were first loaded and then re-sized to get all consistent sizes. Because the images were displayed in BGR format, they were required to be converted to RGB format so that they could be read by the *OpenCV* function.

The input images were down sampled to 224 x 224 or 128 x 128 to reduce their dimension. One obvious reason to downsize the image was that a larger size would require huge memory usage and huge network parameters. In addition, the filter size was also needed to be doubled to have the same receptive field. Reducing the batch size and reducing input image resolution would help to reduce GPU's memory usage when running out of memory occurred.

3.3. Model Development

There were two phases of model development: the training phase and the testing phase, as shown in figure 2. Before the network was trained, all images of the flowers needed to be stored in the image database. After that, all images in the database were pre-processed using several techniques of image pre-processing before it could be used to train the network engine. After training the neural network, the

network was used for the classification process. In this study, two CNN models were compared: *AlexNet*, and *VGG16*.

3.3.1 AlexNet. Five convolutional layers and three fully linked layers make up the network [25]. *AlexNet* achieves transfer learning on a pre-trained network based on the *ImageNet* benchmark, which is fine-tuned in the last three fully connected layers. The network's large number of layers has a good influence on feature extraction. Furthermore, the number of factors has a detrimental impact on performance. *AlexNet's* architecture is depicted in Figure 4. The convolutional layer is *AlexNet's* initial layer followed by the pooling and normalization layers. The *Softmax* is the final layer, and it is this layer that does the classification process.

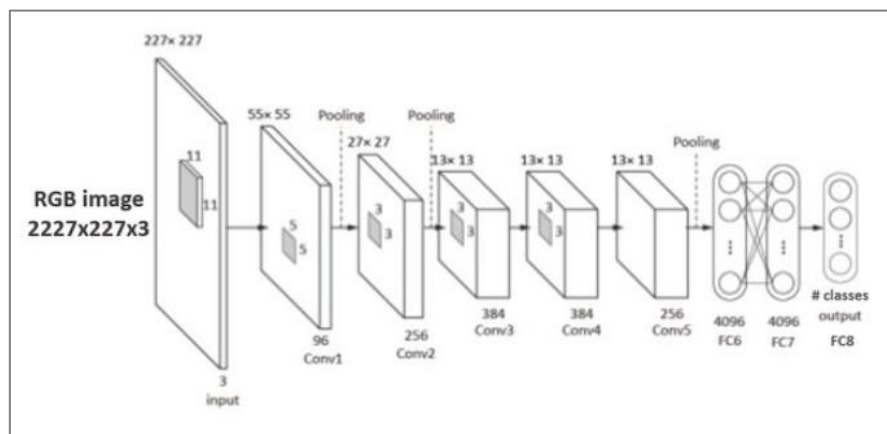


Figure 4. *AlexNet* architecture [25].

3.3.2 VGG16. By replacing huge kernel-sized filters (11 and 5 in the first and second convolutional layers, respectively) with multiple 3 x 3 kernel-sized filters one after the other, the model was introduced as an enhancement over AlexNet [26]. *VGG16* has 138 million parameters in total. *VGG16* had been training for weeks on *NVIDIA Titan Black GPUs*. *VGG16's* architecture is depicted in Figure 5. *VGG16* has addressed most of important aspects of CNN architectural design.

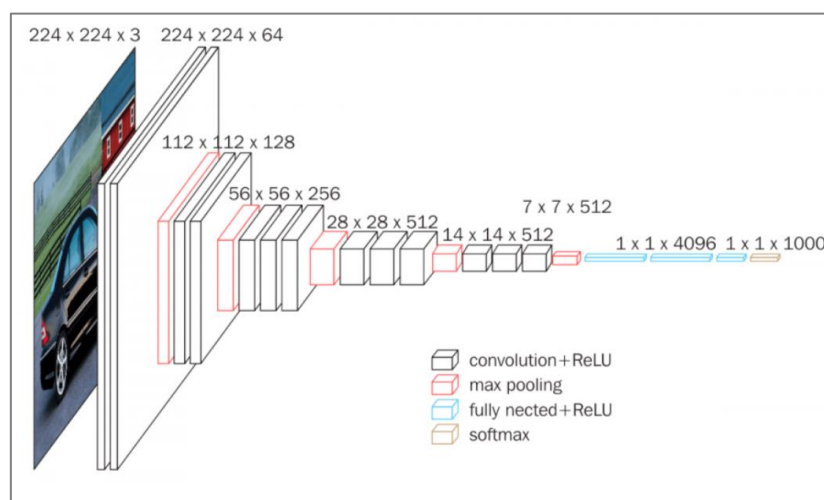


Figure 5. *VGG16* architecture [26].

After each convolution layer, the *ReLU* activation function was employed in the structure where one of the data increasing methods was used. Gradient descent was utilized to train the stack, and the small 3 x 3 filters were used in all layers. A stored network that has been previously trained on a big dataset, generally on a large-scale image classification problem, is known as a pre-trained model. In *VGG16*, the pre-trained model scored 92.7% test accuracy in *ImageNet*, a dataset with over 14 million images belonging to 1000 classes [22].

4. Results and Discussion

4.1. Comparison of accuracy for AlexNet and VGG16

In this study, the pre-trained models were tested by putting them through a classification exercise. A graphics processing unit (GPU) was used to run each model with a maximum of 1,812 iterations. A number of 1,812 iterations was the best iteration value because if the iteration exceeded 1,812 iterations the training loss would decrease, and it could cause the model to overfit the dataset. Figure 6 and 7 show the accuracy graphs for both models *AlexNet* and *VGG16* respectively. Graphs were plotted using *Python* plotting based on the trained image from the datasets.

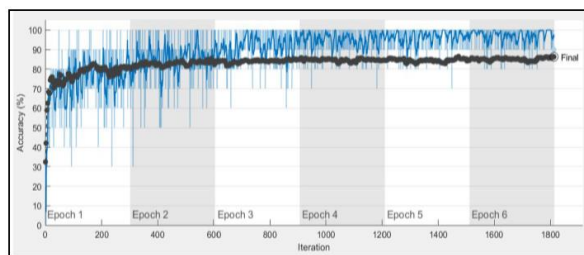


Figure 6. Accuracy graphic for *AlexNet*.

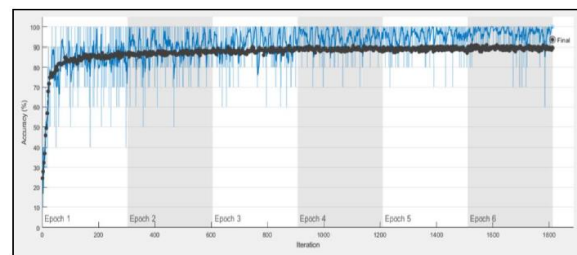


Figure 7. Accuracy graphic for *VGG16*.

The operating performances of the models were tested against *Nvidia GTX1080 GPU* with max iteration 1,812, as shown in table 1. Validation accuracy as a model performance criterion is shown in the table. The *VGG16* network had the highest validation accuracy, according to the findings. Due to the large number of parameters, *VGG16* had the longest working time when it came to operational times.

Table 1. Experimental results.

Models	Validation Accuracy	Time taken
<i>AlexNet</i>	85.69%	200 min 54 sec
	83.44%	302 min 09 sec
<i>VGG16</i>	95.02%	600 min 43 sec
	94.78%	711 min 55 sec

4.2. Final Architecture of VGG16 as Feature Extractor

In pattern recognition and image processing, feature extraction is a type of dimensionality reduction. Feature extraction's primary goal is to extract the most significant information from the original data and express it in a lower-dimensional space [23]. In this study, *VGG16* was used to extract the features based on the results obtained. When a significant amount of data needs to be trained with a GPU, deep learning is used. This is due to the enormous number of iterations or epochs necessary during neural network training as well as image processing, which involves computationally costly 3-d data (i.e. data with width, height, and channels). *VGG16* model architecture contains 16 convolutional layers, 3 fully

connected layers, and 1 Softmax classifier, and uses only 3 x 3 convolutional layers stacked on top of each other. The resulting model of the *VGG16* is shown in figure 8 below.

Layer (type)	Output shape
input_1 (InputLayer)	[224,224,3]
conv1_1	[224,224,64]
conv1_2	[224,224,64]
tf_op_layer_pool1_1	[112, 112,64]
conv2_1	[112,112,128]
conv2_2	[112,112,128]
tf_op_layer_pool2_1	[56,56,128]
conv3_1	[56,56,256]
conv3_2	[56,56,256]
conv3_3	[56,56,256]
conv3_4	[56,56,256]
tf_op_layer_pool3_1	[28,28,256]
conv4_1	[28,28,512]
conv4_2	[28,28,512]
conv4_3	[28,28,512]
conv4_4	[28,28,512]
tf_op_layer_pool4_1	[14,14,512]
conv5_1	[14,14,512]
conv5_2	[14,14,512]
conv5_3	[14,14,512]
conv5_4	[14,14,512]
tf_op_layer_pool5_1	[7,7,512]
flatten	[1,1,25088]
fc6	[1,1,256]
fc7	[1,1,128]
fc8	[1,1,5]
tf_op_layer_Softmax	[5]

Figure 8. *VGG16* model summary.

The input to conv1 layer is of fixed size 224 x 224 RGB image. The first and second convolutional layers (conv1_1 and conv1_2) comprise of 64 kernel filters and the size of the filter is 3 x 3. The dimensions of the input picture (RGB image with depth 3) change to 224 x 224 x 64 as it travels through the first and second convolutional layers. The output is shrunk to 112 x 112 x 64 and passed to the max-pooling layer with a stride of 2. Convolutional layers 3 and 4 (conv2_1 and conv2_2) are made up of 128 feature kernel filters with a 3 x 3 filter size. Following these two layers is a max-pooling layer with stride 2, which reduces the output to 56 x 56 x 128. Convolutional layers with kernel size 3 x 3 are the fifth, sixth, seventh, and eighth layers, which are conv3_1, conv3_2, conv3_3, and conv3_4 accordingly. 256 feature maps are used in all three. Following these layers is a max-pooling layer with a stride of 1. The three fully connected layers which are fc6, fc7, and fc8 follow a stack of convolutional layers: the first layer of fully-connected (fc6) has 256 channels, the second layer of fully-connected (fc7) has 128 channels and the third layer of fully-connected (fc8) performs classification and thus the final classification produces five classes. The model can then be used directly to classify a flower image into one of the five classes.

4.3. Deployment of the best VGG16 model

This section shows the prediction result for random images in the database by using the *VGG16* model. The result as shown in figure 9 shows the predicted flower images from its actual images. The corresponding confusion matrix of *VGG16* is as shown in figure 10.

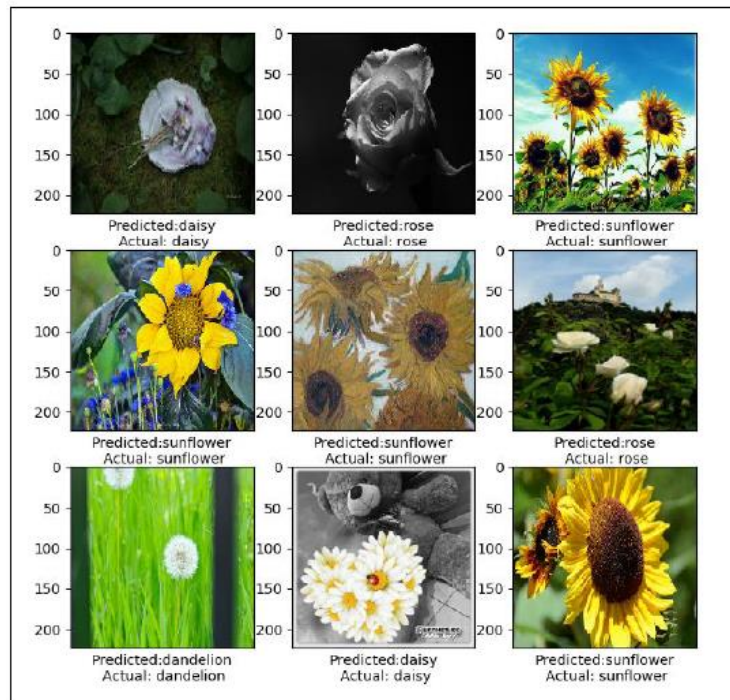


Figure 9. Prediction result for *VGG16*

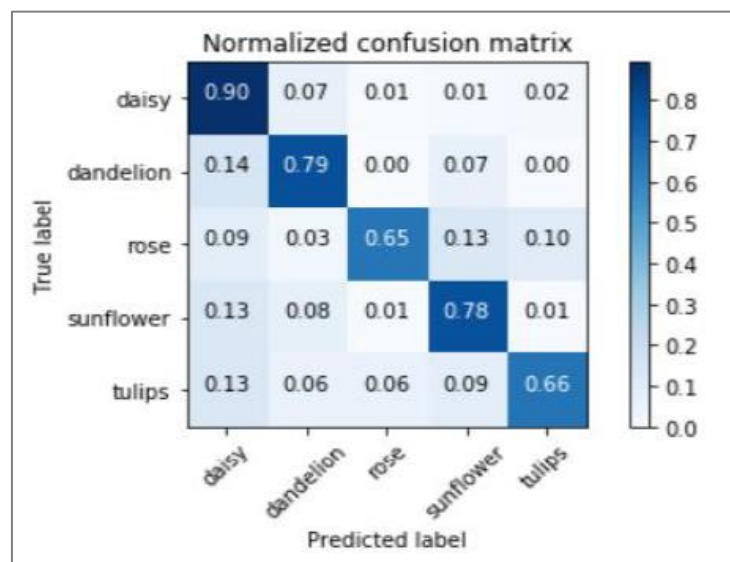


Figure 10. Normalized confusion matrix.

According to the normalized confusion matrix in figure 10, the *VGG16* model shows a flawless performance in predicting daisy flowers with 0.90 normalized value, while dandelion, sunflower, rose and tulip show medium performance with normalized values between 0.65 to 0.79. As mentioned

previously, inter-class similarity and intra-class variation are among the challenges in flower recognition task. Figure 11 and 12 show a few samples of misclassified flowers in different species.



Figure 11. Inter-class similarity on dandelion and sunflower.



Figure 12. Intra-class variability on rose and tulip.

Figure 11 shows the flowers with inter-class similarity (dandelion and sunflower) in which dandelion and sunflower are similar in terms of shape and appearance, while figure 12 shows rose and tulip in which they are identical in terms of illumination and view difference. In this study, dandelion and sunflower whose normalized values are 0.79 and 0.78 respectively show almost identical performance on predicting both flowers because of its intra-class flower similarity. Rose and tulip, with the normalized values of 0.65 and 0.66 respectively, show almost similar performance because of its intra-class flower variability. Based on the result shown in the confusion matrix, it is shown that *VGG16* model network can be used in recognizing the flower images, with excellent performance for distinct species and medium performance for species with inter-class similarity and intra-class variability.

5. Conclusion

This study has shown that flower recognition with deep CNN and with transfer learning approach achieved higher accuracy with *VGG16* as compared to *AlexNet*. Feature extraction with *VGG16* shows a higher degree of prediction on certain flower species, however, flowers with inter-class similarity and intra-class variation show medium prediction performance. This study was tested using the published dataset from Kaggle. We intended to produce a recognition engine for wildflowers found in National Park, which can be done in the next future possible expedition with original samples. A better algorithm should be utilized to handle the problem of intra-class differences and inter-class similarities in flower photos, and this should be combined with in-depth learning expertise to increase flower classification and recognition accuracy.

Acknowledgment

The authors would like to thank Pejabat Timbalan Naib Canselor (Penyelidikan & Inovasi) for their assistance and support in sponsoring our teams in the 2nd Scientific Expedition Taman Negara Pahang, Institute for Biodiversity and Sustainable Development (IBSD) with research grant 600-RMC/DANA 5/3/MITRA SDG (1/2019)-29, and also to the Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia, for the research support.

References

- [1] Mutalib S, Abdullah M H, Abdul-Rahman S, and Aziz Z A 2016 A brief study on paddy applications with image processing and proposed architecture *In 2016 IEEE Conference on Systems, Process and Control (ICSPC)* pp 124-129
- [2] Kahar M A, Mutalib S, and Abdul-Rahman, S 2015 Early detection and classification of paddy diseases with neural networks and fuzzy logic *In Proceedings of the 17th International Conference on Mathematical and Computational Methods in Science and Engineering*,

- (MACMESE) pp 248-257
- [3] Pushpanathan K, Hanafi M, Mashohor S, Fazlil Ilahi WF 2021 Machine learning in medicinal plants recognition: a review *Artif Intell Rev.* **54** pp 305–27
 - [4] Liu W, Rao Y, Fan B, Song J and Wang Q 2017 Flower classification using fusion descriptor and SVM *International Smart Cities Conference* pp 1–4
 - [5] Nilsback ME and Zisserman A 2012 Delving deeper into the whorl of flower segmentation (*Image and Vision Computing*) **28** no 6 pp 1049–62
 - [6] Almogdady H, Manaseer S and Hiary H 2018 A Flower Recognition System Based On Image Processing And Neural Networks *International Journal Of Scientific & Technology Research* **7** pp 166–73
 - [7] Nilsback M and Zisserman A 2009 An automatic visual florasegmentation and classification of flower images *Oxford University*
 - [8] Saitoh T and Kaneko T 2000 Automatic Recognition of Wild Flowers *Proceedings 15th International Conference on Pattern Recognition* **2** pp 507–10
 - [9] Hyeon-Joong Y 2015 Deep Convolution Neural Networks in Computer Vision *IEIE Transactions on Smart Processing & Computing* pp 35–43
 - [10] Shamsuddin M R, Abdul-Rahman S and Mohamed A 2015 Towards applying Deep Neural Network for Complex Input Patterns *Recent Advances in Mathematical and Computational Methods* pp 238–267
 - [11] Taman Negara in Malaysia: The Complete Guide, Visiting Malaysia's Oldest National Park, <https://www.tripsavvy.com/taman-negara-national-park-in-malaysia-4587186>
 - [12] Taman Negara - One of the World's Oldest Tropical Rainforest; <https://www.adventoro.com/blogs/adventure-news/taman-negara-one-of-the-worlds-oldesttropical-rainforest>, Published: May 12, 2017
 - [13] Suratman, M N 2012 Tree species diversity and forest stand structure of Pahang National Park, Malaysia *Biodiversity enrichment in a diverse world* pp 473–492
 - [14] Gogul I and Kumar VS 2017 Flower species recognition system using convolution neural networks and transfer learning 2017 *Fourth International Conference on Signal Processing, Communication and Networking (ICSCN)* pp 1–6
 - [15] Haralick RM, Shanmugam K and Dinstein I 1973 Textural Features for Image Classification, in *IEEE Transactions on Systems, Man, and Cybernetics* **SMC-3** no 6 pp 610–21
 - [16] Löfstedt T, Brynolfsson P, Asklund T, Nyholm T and Garpebring A 2019 Gray-level invariant Haralick texture features *PLoS ONE*
 - [17] Hu MK 1996 Visual pattern recognition by momentos invariants *IRE Transactions on Information Theory* **8** no 1 pp 179–87
 - [18] Kim J, Kim BS and Savarese S 2012 Comparing Image Classification Methods: K-Nearest-Neighbor and Support-Vector-Machines *Proceedings of the 6th WSEAS international conference on Computer Engineering and Applications* pp 133–138
 - [19] Coskun N and Yildirim T 2003 The effects of training algorithms in MLP network on image classification *Proceedings of the International Joint Conference on Neural Networks* **2** pp 1223–6
 - [20] Chollet F 2017 Xception: Deep Learning with Depthwise Separable Convolutions *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* pp 1800–7
 - [21] Wu Y, Qin X, Pan Y and Yuan C 2018 Convolution neural network based transfer learning for classification of flowers *IEEE 3rd International Conference on Signal and Image Processing* pp 562–6
 - [22] Sermanet P, Eigen D, Zhang X, Mathieu M, Fergus R and LeCun Y 2014 Overfeat: Integrated recognition, localization and detection using convolutional networks *International Conference on Learning Representations (ICLR)*
 - [23] Kumar G and Bhatia PK 2014 A Detailed Review of Feature Extraction in Image Processing Systems *Fourth International Conference on Advanced Computing & Communication*

Technologies

- [24] Kaymak C and Uçar A 2019 A brief survey and an application of semantic image segmentation for autonomous driving *Handbook of Deep Learning Applications* pp 161-200
- [25] Krizhevsky A, Sutskever I and Hinton GE 2012 Imagenet classification with deep convolutional neural networks *Advances in neural information processing systems* pp 1097–105
- [26] Simonyan K and Zisserman A 2015 Very deep convolutional networks for large-scale image recognition *International Conference on Learning Representations (ICLR)*